Available online at www.sciencedirect.com

ScienceDirect

Russian Journal of Economics 1 (2015) 81–107

Russian Journal of Economics

www.rujec.org

Behavioral economics and the 'new' paternalism ☆

Rostislav Kapeliushnikov [a,b]

[a] *Institute of World Economy and International Relations, Russian Academy of Sciences, Moscow, Russia*
[b] *National Research University Higher School of Economics, Moscow, Russia*

Abstract

The paper provides a critical appraisal of the normative program of behavioral economics known as 'new paternalism'. First, it explores the theoretical foundations of behavioral economics, describes major behavioral anomalies associated with bounded rationality of economic agents and discusses its normative principles and political implications. It then discusses the main empirical and conceptual drawbacks of new paternalism and provides arguments for the alternative non-welfarist normative tradition based on the idea of freedom.

JEL classification: A13, D03, D60, D63.
Keywords: behavioral economics, welfare economics, new paternalism, normative analysis.

I am a family.
Like a spectrum, seven egos live within me,
unbearable, like seven beasts.
And the bluest blows a shepherd's pipe!
And in the springtime,
I dream that I am the eighth!

Andrei Voznesensky, "Antiworlds and The fifth ace", 1967

1. Introduction

The emergence of a new research program called 'behavioral economics' is one of the most significant events in the evolution of modern economic theory in recent past decades. The program asserted itself in the 1970s as an independent sub-discipline born on the crossroads of economics and psychology. From the very beginning, behavioral economics has challenged the standard (neoclassical) economic theory. From a methodological point of view, its main distinctive feature is the wide application of experimental methods (predominantly in laboratory settings and, to a much lesser degree, in the field). From the point

☆ The updated English version of the article published in Russian in *Voprosy Ekonomiki*, 2013, No. 9, pp. 66-90; No. 10, pp. 28-46.

E-mail address: rostis@hse.ru.

Peer review under responsibility of Voprosy Ekonomiki.

http://dx.doi.org/10.1016/j.ruje.2015.05.004

of view of positive analysis, its most important message is the rejection of the commonly accepted model of rational choice that has traditionally provided the behavioral foundation for the greater part of contemporary economic theory. Numerous experiments performed by behavioral economists have demonstrated that real human behavior shares little with neoclassical theory's main protagonist, the hyper-rational Homo oeconomicus, who possesses a well-ordered set of preferences, perfect information and unlimited computational capacities. In the vast variety of situations, the behavior of real economic agents turns out to be boundedly rational in the best-case scenarios and clearly irrational in the worst-case scenarios.

The ideas and approaches offered by behavioral economics relatively quickly got academic recognition, penetrating the mainstream economic theory and leading to a radical transformation in many of its research domains. Today, an overwhelming majority of economists accept these ideas and approaches at least on the general conceptual level, if not in their actual research practices. The obvious success of behavioral economics is easily evidenced by a partial list of fields it has expanded into: consumer choice theory, finance theory, law and economics, macroeconomics, development economics, game theory, and many others.

Behavioral economists went beyond positive analysis and soon proceeded to develop normative recommendations for the government (and other major players, such as corporations or political parties). The normative program built upon the ideas of behavioral economics has been dubbed the 'new' paternalism. It considerably expanded the boundaries of justifiable state intervention in the economy and individual lives compared with what traditional neoclassical economics was willing to accept. In a popular form, this program was presented in the best seller by Richard Thaler and Cass Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (Thaler and Sunstein, 2008). The key concept of the book, with the term itself (*nudge*) appearing in its title, was almost immediately taken up by economists, lawyers and psychologists, and has since been used to denominate both behavioral economics itself and the normative approach to governmental regulation derived from it. The idea of this manifesto directed at the public is quite simple: 'We [the 'new' paternalists] know how to make your life happier' (Leonard, 2008).

The ideas of behavioral economics proved to be extremely attractive for politicians in many countries across various parts of the ideological spectrum. Barack Obama used them extensively in his election campaigns and has often referred to them during his presidency. *Time* magazine called the group of his most trusted advisors a 'behavioral dream team' aiming to transform the country with the help of implications drawn from behavioral research (Grunwald, 2009). The prime minister of the United Kingdom, David Cameron, has also been quite interested in the ideas of the new behavioral science. He set up the Behavioral Insights Team, a special unit to develop behavioral policies, and invited Richard Thaler, the guru of behavioral economics, as an informal advisor (Wintour, 2010). In Cameron's opinion, the behavioral nudge policy is a perfect mechanism 'to persuade citizens to choose what is best for themselves and society' (Basham, 2010).

The purpose of our analysis is to provide a critical evaluation of the normative attitude of behavioral economics; its positive research program is only discussed to the extent necessary to understand the details of the 'nudge' policy. In our opinion, the gradual transformation of a 'welfare state' into a 'paternalistic state' is one of the vital but poorly comprehended trends in the evolution of the modern government regulation system. This new trend, stemming from the ideas of behavioral economics, deserves closer attention.

2. General characteristics

Behavioral economics, which emerged from a symbiosis of economics and psychology, may be considered another example of a powerful methodological trend in the cross-disciplinary social studies known as 'economic imperialism.' That phrase implies invasion of the

methods and concepts of economic science into adjacent social disciplines. This 'colonization' has already swept through (with mixed results) political theory, sociology, history, law, anthropology, criminology, theology and demography. The word *imperialism*, used metaphorically, points directly to the inequality of parties engaging in a relationship. The advocates of 'economic imperialism' (G. Becker, J. Hirshleifer, E. Lazear, and others) admit that other social disciplines do have valuable observations, concepts and analytical tools at their disposal, but claim that only economic science is able to provide a general conceptual framework for understanding various social phenomena.

However, behavioral economics involves a different type of interdisciplinary collaboration; psychology acts as the 'mother country' whereas economic theory is the territory being 'colonized' (Glaeser, 2004). More specifically, economic phenomena are analyzed with the help of concepts and methods developed in psychology. Such an inversion is quite unusual for modern economics, with its 'imperial' ambitions. As A. Alchian once insightfully stated, 'it is not economic science that is imperialistic; it is the model of human behavior underlying it' (quote from Radnitzky and Bernholtz, 1987). However, to demonstrate the empirical inconsistency of this model is precisely the grand purpose of behavioral economists! Naturally, in this case, interdisciplinary synthesis can no longer be based on the assumptions on human behavior provided by conventional (neoclassical) economic theory.

Another difficulty arises here. The interdisciplinary situations prevalent in modern economics and in modern psychology differ radically. A single paradigm has predominated in economics for a long time, that is, one theoretical approach (neoclassical) has been the mainstream. By contrast, psychology has multiple paradigms without any one school prevailing as the mainstream. The multitude of competing research programs naturally leads to a question: which one has been the main supplier of the methods and concepts 'imported' by behavioral economics?

The presence of the word 'behavioral' seemingly points to 'classical' behaviorism, a leading school of 20th century psychology that is associated with the name of J. Watson . However, a closer examination proves such a reference to be incorrect (Angner and Loewenstein, 2007). Behaviorism assumes that only observed, objectively measurable behavior can serve as the subject matter for psychology. Behaviorists consider it unacceptable for true science to refer to any psychic phenomena (notions, wishes, intentions or plans) that are not actually observed. Behavioral economics' methodology is directly the opposite. Its main objective is to study the effect of various mental — and, therefore, not actually observable — inner states experienced by individuals as part of their decision-making. This points to its close affinity with cognitive psychology, developed in direct opposition to 'classical' behaviorism. This is why many view the adjective 'behavioral' as a misnomer and argue instead that 'cognitive economics' would have been a much closer fit (Lambert, 2006. P. 52).

It is important to distinguish behavioral economics from another research program associated with *experimental economics*. Although the difference between them is not fully clear (depending on the nature of the issue in question, one and the same author may act alternatively as a behavioral or an experimental economist), it nonetheless exists; at least, behavioral theorists themselves insist on its existence (Tversky and Kahneman, 1986). These approaches undoubtedly have a great deal in common. Both address decision-making processes, both use experimental methods, and both place the most value on laboratory research results. However, whereas behavioral economics is focused on individual behavior, experimental economics is more concerned with the results of interpersonal interaction. The former is more interested in the cognitive and behavioral constraints on rationality as such, whereas the latter strives to understand how it is possible to overcome these limitations by drawing on various institutional mechanisms (different sets of 'rules of the game'). The tension between these disciplines resulted in the Nobel Prize being awarded to prominent figures in both behavioral and experimental economics in 2002 — psychologist D. Kahneman and economist V. Smith.

The predecessor to the 'new' behavioral economics is the 'old' behavioral economics of the 1950s and 1960s, associated with ideas of H. Simon and J. Katona (Sent, 2004. P. 742). Simon coined the term 'bounded rationality', which denotes the entire spectrum of constraints on human knowledge and computational capabilities that prevent people in the real world from behaving as predicted by neoclassical theory (Simon, 1987). Katona appears to have been the first who started to use the term 'behavioral economics' (Katona, 1951). Nevertheless, although the ideas of the 'old' behavioral economics raised quite a stir (not to mention earning Simon the Nobel Prize in Economics in 1978), they were almost fully neglected by the overwhelming majority of economists and failed to create any type of a new, specific sub-discipline. More surprisingly, in spite of the obvious parallels, the ideas of the 'old' behavioral economics have not had any significant influence on the 'new' behavioral economics (Angner and Loewenstein, 2007). Apart from the terminological similarity, 'new' and 'old' behavioral economic theories are hardly related.

The behavioral turn in economic theory is considered to have been triggered by two papers published by D. Kahneman and A. Tversky (Kahneman and Tversky, 1974, 1979). The papers criticized the 'orthodox' expected utility theory, offering an alternative concept for decision making under uncertainty that was called 'prospect theory'. Of equal importance in promoting behavioral ideas was the work of economist Richard Thaler that appeared at the same time and provided a great deal of empirical evidence on the 'suboptimality' of decisions made by economic agents, such as underestimating opportunity costs, the inability to abstract from sunk costs, and insufficient self-control (Thaler, 1980, 1985).

Thaler, like Kahneman and Tversky, considered his main goal the development of an empirically adequate theory of choice that would be able to describe actually observed decision-making processes. Inspired by their lead, hundreds of economists and psychologists joined the exciting process of 'deconstructing' the 'standard' rational choice model by discovering more and more holes in it. Renowned scholars such as G. Akerlof, D. Ariely, C. Camerer, G. Loewenstein, D. Laibson, T. O'Donohue, M. Rabin, K. Sunstein, A. Shleifer and others are among champions of behavioral economics. The impact of behavioral economics on the entire spectrum of economic research proved to be so profound and multifaceted that some commentators consider it a true revolution in modern economic thought (Costa-Font, 2011).

3. What is wrong with the conventional model of rational choice?

The standard approach accepted in economic theory presupposes perfect rationality of economic agents. Economists mean several things when referring to 'perfect rationality' (Camerer et al., 2003. P. 1214–1215). First, individuals have well-ordered preferences (objectives) and aim at fully satisfying them when making decisions. Second, they do not make mistakes (at least systematically) when calculating the benefits and costs associated with multiple choices. Third, in uncertain situations, they can make probabilistic estimates of potential outcomes using all available information and revise those estimates as soon as new data arrives. Obviously, the first point is the most important.

The concept of rationality, as understood by modern economic theory, is purely formal. It says nothing about the 'rightness' of the goals pursued by individuals. In this formal sense,rationality is a synonym for a consistency of preferences that are manifested in the actual choices made by an individual. According to Arrow, 'the major meaning of rationality is a condition of consistency among choices made from different sets of alternatives' (Arrow, 1996. P. xiii). Consequently, the restrictions imposed by economic theory to consider a person's behavior 'rational' are minimal; the underlying preferences may be of any type except that they may not be mutually exclusive. Only this approach makes consistently maximizing behavior possible. Given a certain order of preferences and a certain set of constraints (physical, institutional, informational), individuals choose the best options of those available to them.

From the viewpoint of observed behavior, first, this implies that in identical situations, rational individuals will make identical choices (choose the same options). Second, whatever choice they make, they will have no reasons to regret them or take them back (Saint-Paul, 2011). Of course, this does not mean that they will behave similarly in any circumstances every time. In choosing between an apple and an orange, a rational agent may choose the apple today and the orange tomorrow. However, if asked whether she regretted her choice yesterday, she would deny this. Moreover, the apple she took yesterday may have been wormy and she would have chosen the orange had she known that. However, if we put her back in yesterday and provided the same limited information that was available to her then, we would observe her choose the apple.

The analysis of the formal requirements to be met by choices (and, correspondingly, the underlying preferences) to guarantee their consistency is the subject of a vast number of studies (see Tversky and Kahneman, 1986). The lists of these requirements (axioms of rational choice) vary for different authors, with two appearing as the most significant in terms of behavioral psychology: transitivity and context-independence (a slightly different term is 'independence of irrelevant alternatives').

The condition of transitivity assumes that if A is preferred over B, and B is preferred over C, then A is preferred over C. Therefore, rational agents can make choices when presented not only with isolated pairs of alternatives but also with multiple options. (An agent with non-transitive preferences is not capable of this: if she prefers apples to oranges, oranges to bananas, and bananas to apples, she will find himself confused if offered all three at once.)

The other condition — context-independence — manifests in many specific ways. In particular, it assumes that a choice between two options is independent from the order in which they are offered. It also assumes that adding a new option to the existing two should not influence the choice unless the third option is preferable to the previous two. Generally, under conditions of context-independence, decisions made by individuals when offered different descriptions of the same problem will remain the same, i.e., the results of their choices do not depend on the form of presentation.

From a more general, philosophical point of view, the premise of rational behavior is equivalent to an assumption that each economic agent has a single self. *Homo oeconomicus*, as imagined by neoclassical theory, cannot suffer from split personality. Otherwise, there would be no consistency or order in her preferences. In other words, she has only one utility function (a single set of preferences), just as each person has a single nose or a single stomach (Saint-Paul, 2011. P. 20).

The neoclassical theory finds both normative and analytical significance in the rationality principle and uses it to evaluate alternative states of the world in terms of better or worse, and for describing and explaining observed economic phenomena. This principle is the starting point for traditional welfare economics, the normative approach of which was called 'welfarist' (Sugden, 2008). This approach interprets preferences as data, whereas the extent of their satisfaction serves as a normative standard in terms of which any individual's welfare is evaluated. In turn, the welfare of a society is understood as an aggregate of the welfare of individuals who make up the society.

It is occasionally asserted as self-evident that individuals know their own interests better than others do (be it other individuals or the state). Reference is also made to the 'consumer sovereignty' principle;in other words, the 'consumer is always right.' A more conceptually sophisticated argument is offered by the theory of revealed preferences; the very act of an individual choosing a given option when some others are available (physically, financially, or informationally)means that it is best for her and is most fully compatible with her preferences because otherwise, her observed choice would have been different.

In behavioral economics, all of the components of the standard rational choice model (normative and positive) came under frontal assault. First, empirical research showed that, in

real life, '[d]eviations of actual behavior from the normative model [of rational choice] are too widespread to be ignored, too systematic to be dismissed as random error, and too fundamental to be accommodated by relaxing the normative system' (Tversky and Kahneman, 1986. P. 252).

Second, the numerous behavioral 'irrationalities' that have been identified undermine the philosophical foundation of the conventional approach, namely the belief that an individual possesses a single decision making center. In cannot be denied that many of the observed anomalies result from the co-existence of several incompatible sets of preferences in an individual's psyche. In fact, there are multiple selves, each of which, when given the right of decision making, makes decisions based on her own limited interests without considering those of the others. Although each of these incarnations can act as a rational utility maximizer, the overlap of their decisions will inevitably result in irrational behavior.

Third, the normative prescriptions of the welfarist approach are being undermined. The evidence gathered by behavioral economists demonstrates in many cases that people have a poor understanding of their own true interests and often act contrary to them. If consumers are capable of acting to their own detriment, the question of why and on what grounds should their 'sovereignty' be respected arises. Problems occur also with the principle of revealed preferences. In this situation, one cannot be certain that the options actually chosen by individuals would always be the best for them out of all of those available and, consequently, that they should have the last say concerning what is good or bad for them. This is why irrational preferences cannot be accepted as a normative standard for evaluating welfare, whether of an individual or of society as a whole.

4. Catalogue of irrationalities

The most general empirical implication of behavioral economic analysis is that people often understand and interpret situations they face in a manner different from that prescribed by the rational choice model. Their decision making may be influenced by information having no actual value; they suffer from excessive self-confidence; they are slow in executing their decisions; they act mechanically; they make mistakes in assessing the probability of future events; and they act on impulse, influenced by quickly changing emotional states. Because of the multitude and variety of cognitive and behavioral errors described in behavioral literature, upon reading it, '[t]he casual reader might have the impression that the rational *homo economicus* has died a sad death and the economics profession has moved on to recognize the true irrationality of humankind' (Levine, 2012. P. 1).

It is very important (and behaviorists insist on this) that behavioral errors by individuals be predictable (Ariely, 2008). In some circumstances, even competent and successful individuals act irrationally, damaging their own long-term interests. Changing the smallest detail in a situation may be sufficient to alter a person's behavior completely. From an analytical point of view, most behavioral anomalies can be interpreted in terms of multiple selves; each has its own scale of preferences, which renders the empirical behavior of an individual inconsistent.

The variety of deviations from the conventional rational choice model can be divided into two large groups: cognitive errors and defects of the will. However, many of them can be considered manifestations of both intellectual limitations and insufficient self-control. The number of cognitive and behavioral errors identified and described by behavioral economists is large and constantly growing. According to one of the (far from complete)lists, the number is approaching 50 (Rizzo and Whitman, 2009. P. 951). Our analysis is limited to a small sample of the most important ones.

Hyperbolic discounting. The traditional economic analysis of intertemporal choice builds upon the fact that individuals prefer present goods to future ones and are willing to sacrifice a larger amount of the latter to gain a smaller amount of the former. The proportions of this 'exchange' are determined by the subjective parameters of time preference, which may vary

greatly between individuals. Some may be very patient (low discount rate), whereas others may be very impatient (high discount rate).

However, a person's discount rate must remain unchanged for long-term decisions to be rational. In other words, the proportion of 'exchange' between any two periods must be defined only by the distance between them and must be independent of the distance between them and the present moment. In this case, for an individual having a discount factor of 0.9, $100 gained in two years' time will be equivalent to $90 gained in a year, just as $100 gained in a year will be equivalent to $90 gained now. Such a discounting algorithm is called exponential becauseas the time distance from the present moment increases, the decrease in the value of future goods follows an exponential curve.

However, behaviorists have found that in real life many (maybe even most) people act as inconsistent 'discounters,' leaning toward hyperbolic, rather than exponential, discounting (Laibson, 1997). The discount rates people use are not static but rather increase as the compared periods come closer to the present moment. Thus, an individual may value $100 gained in two years to be equivalent to $90 gained in a year, and at the same time value $100 gained in a year to be equivalent to only $80 gained right now. In the first case, the discount factor is 0.9; in the second, it is 0.8.

Decisions made using hyperbolic discounting are time-inconsistent. For example, at t_0, an individual with the time preferences described above would be willing to save $85 at t_1 by depositing it with a bank at an interest rate that effectively yields $100 at t_2 because this decision would be attractive given his discount factor of 0.9. However, when t_1 arrives, he will abandon his original decision and spend the $85 instead of saving it because depositing the money with a bank at the same interest rate would no longer be attractive for him at a discount factor of 0.8. As a result, due to hyperbolic discounting, an individual may become torn between opposite decisions, unable to choose between them. He may plan to start saving for retirement or take up a strict diet next year, but will abandon these ideas when the next year arrives.

Behavioral economists consider this behavior obviously irrational. Most often, it is caused by insufficient self-control and reflects some individuals' inability to resist temptations in the present moment. Excessive impatience (high short-term discount rate) will motivate them to make decisions with immediate benefits and delayed costs. Hyperbolic discounting may be the reason for phenomena such as addictive behavior (for example, to drugs or food), regular postponement of important decisions, building portfolios containing incompatible financial instruments (e.g., using high-interest credit cards while at the same time buying low-interest securities), excessive borrowing, or low saving rates.

Hyperbolic discounting may be considered evidence of conflict between the two selves, each with its own special utility function: one impatient and only concerned with the present, the other prudent and future-oriented. The conflict leads to inconsistent behavior; when the right of decision-making is granted to the impatient self, it cancels prior decisions by the prudent self.

'Cold' and 'hot' psychological states. The current emotional state of an individual may have a decisive influence on her choices. In biologically 'hot' states (e.g., anger, fear, admiration, and excitement), individuals are inclined to make spontaneous decisions, whereas in 'cold' states (e.g., calm, composure, and rational thought), decisions are well-measured (Camerer et al., 2003). In other words, in a 'hot' state, people may react suboptimally, overestimating short-term benefits and underestimating the corresponding long-term costs of a given decision. They may become married in a fit of passion or commit suicide in a profound depression. After inhaling the aroma of leather upholstery in an expensive car, a person may be unable to resist the temptation to buy it.

Here, again, we face the conflict between the two selves, 'hot' and 'cold', with different preference scales. It is obviously dangerous to trust the 'hot' self in making vital decisions that the 'cold' self may later bitterly regret.

Errors of optimism and pessimism. The error of optimism makes individuals overly self-confident in decision making. It is an underestimation of the probability of undesirable events capable of inflicting serious, if not irreparable damage. (Thus, drivers generally underestimate their own chances of having a car accident). This error results in excessive risk-taking (e.g., investing in high-risk financial assets). An opposite type of error also occurs, making people lose self-confidence and exaggerate the probability of undesirable events. This leads to unreasonably strong risk aversion.

Availability bias. This bias is essentially the tendency to overestimate probabilities of particular events when an individual directly participates in them or witnesses those events. Thus, having been robbed in a certain part of a city, one tends to consider that neighborhood dangerous although it may not be any more dangerous than other areas in terms of crime statistics. Availability bias can make individuals abandon even projects with very moderate and reasonable risk levels, thereby damaging their own welfare. (In this sense, their actions are the opposite of those resulting from the error of optimism.)

Context-Dependence. Context-dependence can take two basic forms. First, there is the effect of framing, making an individual choice dependent on insignificant aspects of the situation in which it is made. In this case, we are addressing a direct violation of the 'independence of irrelevant alternatives' principle (see above). These cases are situations in which the outcome of a decision is determined by the formal characteristics of its frame, rather than by its contents. For example, people tend to choose *A* if options are ordered *A–B* and *B* if the order is *B–A*. (A textbook example from medical practice is, when told that the survival rate of a given course of treatment is 90%, most individuals agree to undergo it. Conversely, when told that the given treatment has a 10% fatality rate, most tend to refuse it.)

Second, there is an assessment of available alternatives based on comparisons with a certain reference level. The reference standard for comparison may be determined either by the individual's past experience or by the context of decision-making. Thus, an individual may evaluate her welfare based on both absolute and relative income. That is, the higher the income of her reference group, the lower the perceived value of her own income, and vice versa.

The *status quo bias* is another important case of this sort of dependence, using the current state of affairs as the starting point. This bias has been extensively analyzed. It is manifested in an individual's tendency to avoid anything new, although it promises considerable benefits or low costs to abandon the old. One possible reason is the inclination to feel deeper regret about actions rather than about inaction. Another reason is the tendency to procrastinate, i.e., to systematically postpone making important decisions. A third reason is loss aversion in the face of equal gains and losses, i.e., the propensity to assign relatively higher negative value to the latter than positive value to the former. Loss aversion occurs when people evaluate gains and losses in comparison with the status quo, e.g., their current income. An individual whose income has risen from $8,000 to $9,000 may feel much happier than one whose income has fallen from $20,000 to $10,000.

This phenomenon is associated with the famous endowment effect, i.e., a sharp increase in the value of an item right after its acquisition. A fair amount of experimental data demonstrates that the perceived value of an item for people is much higher when they possess it than when they do not. The sum of money they ask in exchange for giving away an item (when they possess it) tends to far exceed the purchase price they are willing to pay (when they do not possess it) (Knetsch, 1989).

The most widely discussed practical example of status quo bias is associated with the choice of the so-called 'default rules.' It is a choice between alternative contract options, one of which is proposed to an agent 'by default.' The agent, however, may refuse it and choose another (Sunstein and Thaler, 2003b). To illustrate this, we can refer to the cases involving the most common type of retirement plan in the U. S., the 401(k) plan (named after the associated article in the U.S. Internal Revenue Code).

U.S. tax laws allow employees to deposit part of their salary to their personal retirement savings accounts prior to withholding tax, according to retirement plans established by their employers. Employees can sign up for a 401(k) plan in a number of different ways. Some companies enlist their employees by default, reserving the latter's right to withdraw. Others do the opposite. They do not enlist their employees by default, although the latter are free to sign up at any time. In the first case, deductions immediately begin to be made from an employee's salary to her retirement account, and she must submit a special application if she wants to stop it. In the second case, deductions from the employee's salary to her retirement account will not commence until she submits a written request. From the standpoint of the standard rational choice model, the above procedures are identical; hence, the proportion of participating and non-participating employees should be roughly the same in both cases. However, in reality this is not the case. A study of data from three large U.S. companies has shown that, after switching to default enlisting, the portion of employees enrolled in the 401(k) plan had grown from 26–43% to 86–96% in half a year (Choi et al., 2002).

Context-dependence may be considered the most obvious violation of the rationality principle and the most salient example of *preferences endogeneity*. It points to an actual lack of preferences rather than their inconsistency. Sunstein and Thaler noted that when decisions are context-dependent, the very term 'preference' loses most of its meaning. In this type of situation, preferences are shaped at the very moment of choice, rather than preceding it. 'If the arrangement of alternatives has a significant effect on the selections customers make, then their true 'preferences' do not formally exist' (Sunstein and Thaler, 2003b. P. 1164).

Even this short and incomplete overview is sufficient to demonstrate how heavy and varying losses caused by behavioral errors may be for the entire society, and for an individual. 'When consumers make errors, it is as if they are imposing externalities on themselves because the decisions they make do not accurately reflect the benefits they derive' (Camerer et al., 2003. P. 1221).

The question of whether the state is capable of eliminating the ensuing welfare losses arises. Behaviorists answer this in the affirmative, offering a wide range of government interventions that can help people act more rationally, based on their own 'true' interests. Behaviorists themselves take up the role of social therapists, prescribing an appropriate course of treatment for the society (Loewenstein and Haisley, 2006).

5. Under the banner of paternalism

Behavioral economists characterize the general normative standard underlying their political recommendations as 'paternalistic.' In social philosophy, it is understood as any form of the third-party interference in a person's life against her will (e.g., by the state, family, or church) because such interference would improve her welfare and/or would not let her harm herself[1]. In other words, paternalism involves acts of coercion (restricting freedom of choice) aimed at improving the welfare of targeted individuals. An individual is assumed incapable of identifying her own true interests, whereas someone who knows better can and must define these interests for her.

Although paternalism can be practiced both by the state and private entities, 'in modern usage, the term usually refers to those laws and public policies which restrict the freedom of people so that their interests may be better served' (Weale, 1991; cited in: Klein, 2004. P. 261–262). A classic example of paternalistic relationships is the parent-child relationship, with the former restricting the behavior of the latter in the latter's best interests with all types of taboos and prescriptions (e.g., do your homework, come home before dark, do not open the door to strangers). In the case of state paternalism, the state performs the function of the

[1] http://plato.stanford.edu/archives/sum2009/entries/paternalism/.

'father' and its 'fatherly care' is performed on individual adults, although they are recognized as capable and responsible citizens by the law.

Paternalism may take on 'harder' and 'softer' forms, depending on how deeply it interferes with individual decisions. Hard paternalism involves heavy restriction of choice opportunities, whereas the restrictions imposed by a soft paternalism are barely noticeable.

Economic theory has usually opposed paternalistic policy, which limits individual freedom of choice. This consistently anti-paternalistic attitude is quite understandable and can be easily explained. Indeed, if 'the consumer is always right' (always rational), then there is simply no room left for any improvement in her welfare through intervention into her decision making. In this case, any government interference can only cause damage by replacing consumers' optimal decisions by different, suboptimal ones.

Consequently, traditional welfare economics allows restrictions on an individual's freedom of choice subject only to the condition that the restrictions are aimed at improving the welfare of other individuals. This implies two theoretical arguments in favor of government intervention, i.e., reference to 'market failures' and redistributive considerations. An unregulated market either does not function sufficiently efficiently (i.e., does not ensure efficient resource allocation) or creates an income distribution structure inconsistent with the views of equality prevalent in a given society. Government intervention is only justifiable in such cases. Direct intervention into an individual's decisions is deemed unacceptable because it lacks any theoretical basis whatsoever.

Traditional regulatory means (e.g., taxes, subsidies and transfers) may change the budget constraints for consumers, but no encroachment upon their 'sovereignty' occurs. Consumers still have the final say; they are given no prescriptions and may decide for themselves how to react to such changes made by the state. In this sense, the traditional anti-paternalistic attitude of economic theory effectively limits potential state expansion.

However, if the ideas of behavioral economics are correct and economic agents possess only bounded rationality, then the situation changes drastically. The usual list of 'market failures' is supplemented with a new, 'behavioral' one. Consequently, the taboo on paternalistic intervention prescribed by traditional economic theory is no longer in effect, and the state gains access to a much broader field of activities. Because people are not insured against systematic errors, a caring government can (and must!) extend a helping hand to steer their behavior along a rational course using any means available.

Undoubtedly, the 'new' paternalism, based on the ideas of behavioral economics, differs greatly from the old paternalism in its normative standard and recommended types of state intervention. The 'old' paternalism was most often tinged with a distinctive religious or moralistic hue. It ignored the preferences and interests of the 'subjects,' literally replacing them with those of the 'caregiver.' It was assumed, explicitly or not, that a paternalistic state knows better than individuals themselves do what their true welfare is.

The 'new' paternalism takes an entirely different stance. The subjective preferences of individuals serve as the normative standard. The 'new' paternalism, unlike the 'old,' is aimed at helping people achieve what they want. In other words, it tries to help them increase their level of subjective welfare, something they cannot do themselves due to cognitive or behavioral limitations.

Another difference of equal importance is that, whereas the 'old' paternalism tried to improve the welfare of individuals by taking away their freedom of choice, the 'new' paternalism claims that government regulation can improve on the individual welfare without (or nearly without) restricting their personal autonomy (Mitchell, 2005. P. 1245). There are differing, although very similar versions of this idea: the 'asymmetrical paternalism' of Camerer et al.; the 'soft paternalism' of Loewenstein and Haisley; and the 'libertarian paternalism' of Sunstein and Thaler.

Asymmetric paternalism is realized through government interventions that create 'large benefits for the people who are boundedly rational …while imposing little or no harm on those

who are fully rational' (Camerer et al., 2003. P. 1219). The proclaimed goal of soft paternalism is 'to enhance decision making without restricting it'(Loewenstein and Haisley, 2006. P. 6).'Libertarian paternalism' strives to maximally preserve the freedom of choice while pursuing a paternalistic policy (hence the term 'libertarian'). The state should only structure the field of choice, leaving final decisions to individuals (Sunstein and Thaler, 2003a).

In spite of these stipulations, 'new' paternalists firmly deny the traditional anti-paternalistic attitude of economic theory. Sunstein and Thaler are the most uncompromising. They believe the 'dogmatic paternalism' of standard economic theory to be founded on one false assumption and two misconceptions. The false assumption is formulated as follows: '[A]lmost all people, almost all of the time, make choices that are in their best interest or at the very least are better, by their own lights, than the choices that would be made by third parties' (Sunstein and Thaler, 2003b. P. 1163). The assumption is invalid from an empirical standpoint and has been disproved by multiple studies performed by behavioral economics (see above). The two misconceptions are that, on the one hand, 'there are viable alternatives to paternalism' and, on the other hand, 'paternalism always involves coercion' (Sunstein and Thaler, 2003b. P. 1164–1165). Sunstein and Thaler believe both to be erroneous.

They base their counter-arguments on a case that became a textbook example of 'new' paternalism. It opens their book, *Nudge* (Thaler and Sunstein, 2008). The example addresses a dilemma facing the director of a cafeteria owned by a firm. Her customers move along the counter, choosing which of the various dishes they prefer. The director noticed that the dishes that were placed earlier in line were in greater demand than were the others (the framing effect). She also knows about medical studies proving that people can improve their welfare by eating fewer cakes and more fruit. Therefore, how should she place the dishes? Sunstein and Thaler identify four options: place the dishes randomly; place the healthiest (in the director's opinion) at the beginning of the line; place those that would make the customers as obese as possible at the front; or place first those that, in her opinion, would best fit their preferences. However, the last option is only possible if the visitors have exogenous preferences independent from the context (i.e., the order of the dishes). If they have no such preferences (people choose cakes if those are placed first; the same with fruit), this non-paternalistic option falls out and the director must choose between the remaining 'paternalistic' options. As the choice of some frame (order of dishes) is inevitable, the guiding effect of the director's decisions on the visitors' choice is also inescapable. Therefore, she could not avoid nudging even if she wanted to. Thus, if you must nudge visitors by selecting this or that order of dishes as the default, at least it should be the healthiest option.

This leads to the conclusion that there is no real alternative to paternalism. 'The central point, 'Sunstein and Thaler say, 'is that effects on individual choices are often unavoidable. Of course, it is usually good not to block choices, and we do not mean to defend non-libertarian paternalism here. But in an important respect, the anti-paternalistic position is incoherent simply because there is no way to avoid effects on behavior and choices. The task for the committed libertarian is not to avoid such effects but to preserve freedom of choice' (Sunstein and Thaler, 2003b. P. 1182).

Sunstein and Thaler also consider the misconception that paternalism always involves coercion to be erroneous. By putting first fruit and then cakes on the counter, the director nudges visitors to choose dishes that are most in line with their interests. Visitors, however, remain free to choose cakes if they want to. In other words, although paternalistic intervention nudges (with the help of the framing effect) boundedly rational individuals in the direction chosen by the director, it will never influence the behavior of fully rational individuals, who will choose dishes of their liking regardless of the order they appear (that is, framing). Thus, libertarian paternalism fully respects the preferences of rational consumers. On the other hand, because irrational consumers lack any structured preferences, there is nothing to respect. Thus, the nudge policy manages to combine paternalism and libertarianism,

incompatible as these may seem. These arguments lead Sunstein and Thaler to the conclusion that a transition is needed from the old, dogmatic and anti-paternalist attitude that economic theory has traditionally followed to a new, anti-dogmatic and paternalist one (Sunstein and Thaler, 2003b. P. 1161–1162).

Nevertheless, these general declarations are not sufficient for the practical implementation of 'new' paternalist ideas. First, it is necessary to determine which preferences of the multiple selves should be considered reflecting the 'true' preferences of an individual. Whose acts of choice should be taken as the starting point and whose should be discarded? 'New' paternalists try to solve this problem by using the criterion of 'informed desire' (Sugden, 2008. P. 232). This criterion implies that individuals act against their own interests when making decisions 'that they would change if they had complete information, unlimited cognitive abilities and no lack of self-control' (Sunstein and Thaler, 2003b. P. 1162). The preferences of selves whose decisions are the closest to the ideal of full rationality should be taken for granted. The preferences of selves whose decisions deviate from the ideal should be corrected by the state's paternalistic interventions.

This point of view implies that individuals do possess a set of well-structured 'true' preferences that, however, are manifested in a distorted form in observed acts of choice due to various cognitive and behavioral biases. Thus, the task is to reconstruct them somehow. As follows from the above quote by Sunstein and Thaler, the 'new' paternalists suggest solving it in a thought experiment, i.e., imagining what a given individual would choose if he were fully rational. The preferences thus reconstructed are considered the 'true' ones and become the normative standard for evaluating any forms of state intervention aimed at improving the welfare of individuals or that of society as a whole.

As a result, the position of behavioral economists becomes ambivalent with respect to the standard rational choice model; they decisively reject it as a descriptive theory but preserve it and strongly insist on its unconditional significance as a normative ideal. Thus, the state becomes a tool they intend to use to push the empirically observed behavior of boundedly rational individuals closer to the theoretical ideal of full rationality, i.e., convert them from boundedly to unboundedly rational. In this sense, behavioral economists do not abandon the standard model of rationality; rather, they call for practically implementing it as fully as possible. 'The irony is that behavioral economics, having attacked *Homo oeconomicus* as an empirically false description of human choice, now proposes, in the name of paternalism, to enshrine the very same fellow as the image of what people should want to be. Or, more precisely, what paternalists want people to be' (Leonard, 2008. P. 257).

6. Forms of behavioral policies

The normative activity generated by behavioral economics is exceptional. An estimate shows that in only five years (from 2005 to 2009), American law journals alone published approximately 1,000 articles with proposals inspired by behaviorist ideas (Wright and Ginsburg, 2012).

The set of policies supported and approved of by the 'new' paternalists is a mixture of legislative and administrative restrictions, taxes, information disclosure schemes, and other means of convincing and manipulating the 'choice architecture' (a term coined by Sunstein and Thaler to denote alternative ways to describe one and the same choice situation). This involves not only the new, original recommendations but also a multitude of traditional means of state regulation that have been used before, but either without any theoretical support or by using the standard arguments drawn from traditional welfare economics (redistribution considerations or references to 'market failures').

Behaviorists tend to consider any form of government regulation desirable and admissible if they effectively eliminate cognitive and behavioral errors (Rizzo and Whitman, 2009). Not all forms are equal; some are preferred to others (see above).

We will first examine the strongest form of intervention, i.e., direct prohibitions that limit the possibilities of individual choice. These prohibitions, in the opinion of the 'new' paternalists, can and should be imposed when consumer irrationality is too high to be corrected using softer means (e.g., disclosure requirements or 'choice architecture' manipulation). Direct restrictions are considered acceptable in all cases associated with high risk. For example, adherents of behavioral economics support proposals to ban smoking in public places, protect the mandatory government retirement insurance system, prohibit drug use, and so on.

'Sin taxes', such as high excise taxes on alcoholic beverages, tobacco, and gambling, also enjoy support from behavioral economists. Many add another item to the list: junk food, such as fatty products and carbonated drinks. By increasing the direct costs of temptations, sin taxes may help to reduce the negative effects from defects of will and hyperbolic discounting, nudging boundedly rational individuals toward more reasonable behavior.

A rather simple method recommended for correcting mistakes made by people in 'hot' states of mind is to establish so-called 'cool-off' periods in the law. These periods can be added before and after making important decisions (Camerer et al., 2003).

Marital law provides an example of setting such a period ex ante. It usually requires that the official registration of a marriage be performed upon expiry of a certain period rather than immediately after the application is filed. This period is granted to prevent any possible effects from a 'hot' state of mind on such an important decision as creating a family. Examples of ex-post cool-off periods can be found in consumer rights legislation. When purchasing certain items (as a rule, expensive ones), such laws entitle consumers to return the item within a certain period with full reimbursement. Cool-off periods are also needed when making charitable donations (the donor may demand her money back within a certain time), for door-to-door trade, and so on.

Another tool actively supported by the 'new' paternalists is mandatory information disclosures when making any major transaction involving a lease, mortgage, loans, and high-risk goods and services. Thus, they argue that sellers of harmful products (e.g., cigarettes), apart from general hazard warnings, must provide detailed information concerning the respective risks (e.g., with explanations or statistics). To fight the irrational passion for lotteries, behaviorists suggest printing detailed calculations of actual winning chances on lottery tickets.

However, providing abstract information on potential risks to boundedly rational individuals (e.g., those prone to the error of optimism) may not yield desired results. In such cases, 'new' paternalists suggest targeting the subconscious rather than the conscious. To fight the error of optimism, another error can be used that is based upon availability bias. For example, this can be done by printing worrisome phrases on hazardous products (such as 'Smoking kills') and, if that does not help, by adding deterring stories and pictures. If this does not work either, some behaviorists believe consumers can even be told false deterring information. These risk narratives are purported to facilitate decision making under high risk for boundedly rational individuals.

Adherents of behavioral economics most strongly recommend the policy of mandatory disclosure for consumer credit (Wright and Ginsburg, 2012). They do so because behavioral errors are quite common in finance, and the costs are very high. Under direct pressure from the behaviorists, the design of many loan products was changed in the U.S. and the Consumer Financial Protection Bureau was established in the executive branch. Disclosure requirements are only the beginning of behaviorist recommendations. They also demand that mandatory standardized products be used as default options in loan agreements. In their opinion, before imposing complicated, 'fancy' loan options with numerous additional clauses on consumers, banks must first offer the simplest, most basic options. Another actively discussed proposal addresses separating the saving and transactional functions of credit cards. Cards should be 'split' in two, so that one is used only to credit funds, and the other to pay for goods and services. Some behaviorists have voiced the highly radical idea of a full legislative ban

on credit cards altogether because boundedly rational individuals cannot use such a sophisticated financial instrument without harming themselves (Wright and Ginsburg, 2012).

However, the favorite behaviorist idea, of course, is manipulating the 'choice architecture' or, in other words, the soft nudge policy. It is a new innovative contribution from behavioral economics to the practice of state regulation.

We have seen with the retirement plan examples that boundedly rational individuals may make completely opposite decisions depending on which of the available options are offered to them as the default. This is only one case out of many similar ones. Thus, in New Jersey and Pennsylvania, consumers were allowed to choose between partial and full coverage when signing auto insurance contracts. Car owners were offered partial coverage as the default option in the former state and full coverage in the latter. As a result, 20% of owners chose full coverage in New Jersey, whereas75% did so in Pennsylvania (Camerer et al., 2003. P. 1227). Most European countries employ a system that assumes by default that every person agrees to donate his organs after death. If a person objects, he must make it known (a mark is placed on the driver's license). The U.S. used the opposite approach, which assumes by default that a person does not agree to donate his organs to others. If a person chooses otherwise, he must make it known (a special mark is placed on his driver's license). Thus, over 90% of the population are potential donors in Europe, whereas the respective share in the U.S. is under 30% (Sunstein and Thaler, 2003b. P. 1192).

When choosing whether to participate in a retirement savings plan, new paternalists insist that automatic enlisting should be the default option offered to employees because it would help rule out errors associated with the lack of willpower or hyperbolic discounting, for example. A similar approach should be used, in their opinion, when choosing between employment contract options (Sunstein and Thaler, 2003b), e.g., between contracts allowing the dismissal of employees at will, or only subject to objective, strictly defined rules based on a limited set of reasons (dismissal at cause). Experience shows that when the default option is the dismissal at will, most employees agree to it. However, when offered the option for dismissal at cause, most employees select this option. According to the behavioral approach, to increase employee welfare, the second, 'friendlier' approach should always be offered to employees as the standard contract (default option)[2].

However, the nature of paternalistic recommendations concerning the choice of default options may differ considerably. In certain cases, advocates of 'new' paternalism only try to convince firms to use the default options that are most in line with 'true' employee interests when choosing retirement savings plans. In others, they insist that the use of employer-friendly options be fixed in the law. In still other cases, they go even further and call for firms to be allowed to abandon friendly versions of employment contracts only in exchange for hefty employee compensation (Rizzo and Whitman, 2009).

At the same time, behavioral economics defies or at least questions many traditional approaches to government policy, in particular, any schemes that involve regulating and monitoring current economic behavior by using future rewards or penalties.

For instance, in many countries, unemployment insurance systems do not provide benefits to people who decide to leave a job on their own. Thus, they make employees consider the future negative effect (zero income) they create when making current decisions (to quit/not to quit). Threatening a serious punishment tomorrow for irresponsible behavior today, they encourage employees not to make decisions that generate negative externalities for society (in this case, higher expenses for unemployment benefits).

However, the situation is different in the case of the multiple selves identified by behavioral economic analysis. The decision to quit is made by one self, whereas the other is left without

[2] Another illustration of the nudge policy is the (rejected) proposal of New York mayor M. Bloomberg to ban sales of carbonated drinks in large bottles in public entertainment areas — cinemas, parks, and so on. As a result, he said, the consumption of this unhealthy product would be reduced considerably (due to the framing effect).

an unemployment benefit. If the first self is concerned only with her own welfare and is indifferent to the welfare of the other, any attempts to influence her behavior by punishing the other self will have no effect. Denying the benefit will only aggravate the other self's situation without changing anything about today's behavior.

In other words, when addressing boundedly rational individuals, the state should rely not only on incentive schemes associated with future rewards or penalties but also on preventive (including repressive) measures that punish current behavior (Saint-Paul, 2011). In this situation, future selves find themselves in a weaker position to negotiate and should not be left without government support. (In our case, persons quitting their jobs should be equally entitled to the same benefits as those who are fired by an employer.) 'Direct restrictions on current (irrational) selves; unlimited social support to future (rational) selves'—this is how a generalized formula for the behavioral approach might appear.

7. From a Welfare State to a Paternalistic State

Adherents of the 'new' paternalism seek to accentuate in every possible way its moderation. In the opinion of C. Jolls and C. Sunstein, it is somewhere inbetween a laissez-faire system and the traditional ('hard') paternalism (Jolls and Sunstein, 2006). Camerer and his co-authors describe their approach as 'careful, cautious and disciplined' (Camerer et al., 2003). Sunstein and Thaler define their ideas as a 'comparatively weak and non-intrusive type of paternalism' (Sunstein, Thaler, 2003. P. 1162). However, is it really so?

First, the scope of paternalistic recipes, as understood by the adherents of behavioral economics, appears to be extremely broad (potentially, unlimited). For example, according to Thaler and Sunstein, 'people will need nudges for decisions that are difficult and rare, for which they do not get prompt feedback; and when they have trouble translating aspects of a situation into terms that they can easily understand' (Thaler and Sunstein, 2008. P. 74). In other words, they make the 'nudge' policy applicable to practically all possible deals, excluding regularly repeated transactions with ordinary goods and services.

Second, this policy is addressed to the society as a whole and not to particular, specific groups—because in certain situations, even functionally successful individuals can make serious cognitive and behavioral errors.

Third, the range of tools and methods of state intervention, which adherents of the behavioral approach are ready to apply, also appears to be extremely wide. They do not limit themselves to the techniques of the 'soft' paternalism they advocate; when necessary, they are also ready to resort to all of the time-tested methods and tools associated with old-fashioned 'hard' paternalism. Though other conditions being equal, the preferred methods are interventions that minimally abridge individual freedom of choice; direct bans, directive orders and high taxes are not rejected either. When 'soft' paternalism measures do not yield the desired effect, the state may invoke 'non-libertarian' interventions that placed tight constraints on individual freedom of choice. Figuratively speaking, when the 'soft nudge' policy fails, it easily and smoothly turns into a policy of heavy 'kicks and blows.'

Fourth, within the framework of the 'new' (and of the 'old') paternalism, relationships between the state and society develop along the same lines as relationships between adults and children. And this is a source of threat for the growing infantilization of society. Here is the vicious circle: the more infantile a society is, the more it must be paternalized by the state; but the stronger this need, the less responsible and independent members of society become.

Fifth, the 'new' paternalism casts away normative restrictions on state activism that exists within traditional welfare economics. Traditional welfare economics considers state intervention as justifiable only if individual decisions affected some third party. By contrast, behavioral economics justifies extensive government interventions even when an individual's decisions involve no one but himself (when there are no distribution or external effects whatsoever). Consequently, the scope of government control is dramatically extended.

Sixth, the very nature of state intervention changes. It goes beyond changing the external environment, in which individuals make decisions (in sets of incentives, to which they are supposed to respond). A paternalistic state reaches further and starts making decisions for and instead of individuals — either implicitly (by using various 'nudges') or explicitly (by using direct bans and orders).

Seven, most interventions advocated by adherents of the 'nudge' policy manifest themselves not consciously but subconsciously. In some cases, people may not even suspect that they have become an object of manipulative or corrective actions of politicians. As a consequence, such actions by the state spin out of the rational control of individuals.

Finally, in the long term, the 'new' paternalism may lead to the formation of a caste society consisting of two different groups of individuals, rational and irrational, with different rights and different scopes of responsibility. It is evident that irrational individuals may not be granted the same rights enjoyed by rational individuals and may not bear the same responsibility for their actions.

All of these are fundamentally new characteristics that are not typical of the well-known canonical model of a welfare state. From here, G. Saint-Paul makes quite a plausible conclusion that we observe the process of the traditional 'welfare state' transforming into a 'paternalistic state.' In his view, this new paternalistic model rests on four main pillars (Saint-Paul, 2011).

1. To solve their behavioral problems, people need third-party intervention. The best candidate to play this role is the state.

2. Experts can study the effect that certain choices have on real-life welfare and make decisions on behalf of individuals that are better than decisions that will be made by the individuals themselves.

3. Any incentive schemes that make people responsible for the consequences of their past actions are ineffective. These should be replaced with schemes that would immediately reward or punish people for the future consequences of their current actions — consequences that they themselves cannot realize and consider.

4. In terms of politics, how people feel about themselves in society is more important than what they want or do.

Let us analyze each of these points in more detail.

1. Having to address insufficiently rational individuals, a state cannot limit itself to securing only their protection from others. It is also the state's duty to protect them from themselves. Because individuals are not capable of taking care of their 'true' long-term interests, there should be someone else to do that for them, and that role is best suited to the state. First, when developing a paternalistic policy, the state can use expert recommendations; second, for the effective implementation of such a policy, the state has a perfect tool, i.e., the exclusive right of use of coercion. Therefore, the state can make the life of partially rational individuals better by forcing them ('like it or not') to become happier. Private agents and other institutions lack such capabilities and, therefore, are not good candidates for this purpose.

2. To implement a paternalistic policy, the state needs exact and accurate information about the psychological mechanisms governing human behavior. Such information can only be obtained through extensive social and psychological studies. This attributes a critically important role to scientists involved in such studies. Academic science is beginning to have the last word in determining political measures that should be taken by the state. Its assigned task is to identify 'true' individual preferences and then recommend to the state the most effective tools for satisfying them or, in other words, the most effective means of saving people from their cognitive and behavioral errors. In this sense, the key figure is no longer an economist but a psychologist, and society can be compared to a large psychotherapeutic clinic. From this point, it is not society but a narrow group of experts that is supposed to determine what is good and what is bad. This sharply narrows the scope of potential applications for the general constitutional norms, as there are no guarantees that they will always and in every respect agree with the conclusions of the new behavioral science. At the same

time, there is a risk that the results of social and psychological studies will be misused by politicians and special interest groups.

3. Displacement of ex post incentive schemes with ex ante preventive restrictions has far-reaching consequences for the functioning of the legal system, for criminal and civil law in particular. In the context of multiple selves, it appears that one (irrational) individual commits a criminal act but another (rational) individual must pay for it. It is as logical as punishing an innocent person for a crime that he has nothing to do with. Moreover, in the case of abandoning the single self-concept, the very idea of punishment largely becomes senseless. By punishing a rational self, it is impossible to interdict an irrational self from committing a crime, as the irrational self is driven by his short-term interests, with no consideration for long-term negative consequences that may appear and that will not be faced by the current self but rather by another, future self. To avoid such collisions, the emphasis in state policy should be shifted from punitive measures to preventive measures;current selves should be subject to conditions under which they will not be able to behave irrationally (specifically, to commit crimes).

The situation with civil law is similar. Under the condition of multiple selves, the terms of negotiated contracts cannot be considered mutually beneficial by definition. It is one self that signs the contract, but another self that bears obligations under it. In such a situation, the state itself should define the basic contract terms so that the current selves can do no harm to the future selves. In the long term, consolidation of a paternalistic state could lead to reformatting the entire legal foundation of modern societies.

4. Because choices made by partially rational individuals do not reflect their 'true' preferences, they cannot serve as a normative benchmark for state policy. Such 'true' preferences can be reflected by the social feeling of people — to what extent they are satisfied or dissatisfied with life. Accordingly, state policy should be responsive not so much to changes in the economic status of people or in their behavior as to changes in their subjective welfare.

Of course, the exact shape of the future model of a paternalistic state is not yet clear. It is being formed gradually, at a grassroots level, through the progressive build-up of localized interventions, which are frequently not inter-connected and seem insignificant at first sight. A paternalistic state intrinsically grows out of a welfare state and retains many of its characteristic features. In particular, its objectives are identical to those of a traditional welfare state (Saint-Paul, 2011). However, seeking to attain such objectives, it starts using much more 'petty' and intrusive forms of control and regulation.

Surely, it would be naive to believe that a paternalistic state emerges as a direct result of the appearance of behavioral economics and the widespread acceptance of its ideas. It is rather about two parallel processes — real and intellectual. Paternalistic patronage of the society is, apparently, a 'genetic' predisposition for the institution of state as such. However, it would be a mistake to underestimate the role of behavioral economics in this process. Behavioral economics rationalized many forms of intervention, which were spontaneously used by the state, by substantiating them with a theoretical foundation and thus making them intellectually respectable. At the same time, it suggested a great many new, more-refined tools for paternalist policy. It was due to behavioral economics that, as Saint-Paul stated, the last bastion of rationalism in social disciplines fell, such bastion being represented until recently by economic science (Saint-Paul, 2011). In all of these respects, behavioral economics undoubtedly contributed to the gradual shift toward a paternalistic bias in the activities of modern governments.

8. Critical appraisal

Behavioral economics is a complex and heterogeneous phenomenon, and its normative program can be analyzed from many different angles. We will focus on a number of inter-

related questions. How reliable are empirical data underlying its normative conclusions? To what extent are the key tenets feasible, and what obligations do they impose on the state? How convincing is its theoretical justification? Finally, are 'new' paternalism and the traditional 'old-fashioned' paternalism really divided by a tight conceptual boundary, as adherents of the behavioral approach are prone to believe?

1. Though it is already quite mature, behavioral economics still appears to be just a roster of separate empirically observed psychological phenomena. There have been no attempts to join these under any particular synthetic concept. Behavioral economics suggests no 'general theory of cognitive errors;' moreover, its advocates do not seem to need such a theory (Wright and Ginsburg, 2012). As a rule, they content themselves with simple descriptions of certain repeating behavioral reactions without inquiring about their possible causes (Grüne-Yanoff, 2012). Consequently, the political program of 'new' paternalism also lacks any general conceptual foundation and reverts into many small and petty interventions 'on the occasion,' which are almost unrelated to each other.

2. Equally, there have been no systematic attempts within behavioral economics to assess the frequency of various cognitive and behavioral errors. What is the quantitative ratio of rational to irrational individuals? It is one thing if, say, a vast majority of the society is subject to hyperbolic discounting. It is another thing if it applies to a pitiable minority; it is yet another thing if we can talk about hyperbolic discounting only with respect to certain social groups. Usually, behaviorists only make general statements about an extremely non-uniform distribution of behavioral anomalies within a population. However, this selectivity has no effect on their normative recommendations, which are applicable by default to the entire society, without exclusion. This applicability is justified by the ideas of 'asymmetric' or 'libertarian' paternalism. Since, as is asserted, the interventions resting on these ideas impose almost no costs on rational individuals and preserve their freedom of choice, then even if the state acts assuming (incorrectly) that all members of the society are irrational, it will not cause any serious problems.

However, even a minor shift towards measures used with traditional ('hard') paternalism can invalidate these arguments. At this point, empirical data on the distribution of behavioral errors among the population becomes critical. Without considering this distribution, a paternalist policy will most likely fail and is likely to worsen the welfare of the community.

3. The normative program of the so-called 'new' paternalism is built upon a shaky empirical foundation. The experimental findings it appeals to are far from conclusive enough to make firm recommendations for the state. The most striking illustration of this disconnect, is, perhaps, provided by the famous 'endowment effect' first recorded by Thaler (Thaler, 1985), which for many years was regarded as the most reliable and important empirical finding among those ever made in behavioral economics.

Interest to this effect was fueled, largely, by the far-reaching practical conclusions that follow from it. The fact is that the endowment effect is in direct contradiction to the predictions following from the Coase Theorem and can be regarded as its refutation. With this particular effect, the assignment of property rights ceases to be a neutral factor and starts to influence the structure of production and efficiency even where transaction costs are zero. Indeed, if the very fact of possessing a certain good automatically increases its value for individuals, then property rights can remain forever with less effective owners and never pass on to more effective ones. In such a situation, the issuing of the initial delineation of property rights becomes crucial, and the legal system should be developed in consideration of this factor[3].

[3] According to some estimates, out of all the articles published in American law journals over the past two decades, more than one thousand referred to the endowment effect. Its discussion resulted in numerous proposals to completely revise the existing norms in various legal domains, e.g., property, tort, contracts and intellectual property (Wright and Ginsburg, 2012. P. 11).

Laboratory experiments proving the existence of the endowment effect are usually conducted using certain cheap objects. For example, one group of subjects is given mugs, while the other group gets chocolate bars, of an approximately equal value. The mug owners are then offered to exchange their mugs for chocolate bars and, in parallel, chocolate bar owners are offered to exchange their chocolate bars for mugs. The very first experiment brought about surprising results; in the first group 89% of subjects refused to part with their mugs and 90% of subjects in the second group refused to give away their chocolate bars (Knetsch, 1989). From this finding it was concluded that individual preferences depend upon the objects they happen to be 'endowed with' in laboratory conditions; this is why in the first case most of the participants preferred to keep their mugs, whereas in the second the participants chose to keep chocolate bars. The subjects from both groups refused to part with what they came to own during the experiment. This actually implies that when selling an object (owned by them),participants set a much higher price for it than when buying it (when it is not owned by them).If agents had exogenous preferences that were independent of the context (in the case in question, the initial ownership), these results would not be possible. The sale and the purchase price would necessarily be identical.

What is the cause of such an exchange asymmetry? A standard explanation attributes it to the endowment effect, which, in turn, arises from the fact that an individual is subject to loss aversion. According to this explanation, people assess their losses from parting with an item and the benefits from acquiring it differently. The assumed losses almost always seem more significant to them than the assumed benefits. As a result, when selling an item, they set the price higher than what they would agree to as a purchaser.

However, this conventional explanation was challenged in the studies of Ch. Plott and K. Zeiler (Plott and Zeiler, 2007; 2011). They noted that the objects endowed during the experiment may be perceived by subjects as gifts from organizers; the very fact that the object was endowed by the experiment organizers may be interpreted by subjects as a proof of its higher quality; when subjects were offered to sell the object (in the first part of the experiment), the endowed object was right before their eyes, whereas when they were offered to buy the object (in the second part), it was not presented to them; decisions to sell/buy objects were made publicly, in the presence of other participants.

After Plott and Zeiler modified the procedure to take these factors into account, the exchange asymmetry disappeared. The number of participants who preferred to keep mugs (in the first part of experiment) practically equaled the number of participants who chose to buy them in exchange for chocolate bars (in the second part of experiment). In other words, upon a closer analysis, the endowment effect appears to be an artifact. It was vividly demonstrated that the cause of the exchange asymmetry from which it originates is not directly related to 'loss aversion,' and that the exchange asymmetry implies no underlying behavioral dysfunctions, but rather fits into the conventional model of choice quite well.

This case study of the best-known behavioral 'anomaly' illustrates a more general problem concerning any potential manifestations of irrationality, their dependence upon the smallest details of how the experiment is arranged. More specifically, we can never know for certain what is behind a specific assumed anomaly, a real contravention of the rationality principle or some peculiarities of the laboratory procedures. A question arises as to whether state policy be built upon observations that are so sensitive to even minor changes in the setup of laboratory experiments.

4. Adherents of the 'new' paternalism implicitly assume that politicians know the 'true' preferences of individuals better than the individuals themselves. However, the fact that real people are not perfect and can make errors (even systematically) is not enough to conclude that people need or desire to be 'paternalized' by the state. The 'nudge' policy should be analyzed in a comparative institutional perspective.

Individuals can be imperfect utility maximizers and, at the same time, be capable of reaching higher levels of utility when making choices on their own, compared to when choices

are made for them by third parties (including the state). They may not fully realize their own preferences, but still they know more about them than anybody else, even the people closest to them.

This suggestion is supported by the results of a study by J. Waldfogel (Waldfogel, 2005). The study compares how individuals evaluate items that they 'gifted' to themselves during the holidays and gifts that they received from relatives and friends. It appeared that on average the value of the items that individuals bought themselves was approximately 18% higher than the value of gifted things, even though the gifts were made by the closest associates, who should know well the preferences and wishes of the gift recipients.

As is to be expected, friends and relatives should know more about a person's 'true' interests than public officials or politicians. So, if even 'close' associates have only a rough idea of what individuals whom they know very well prefer, what can be expected from 'far' associates — for example, agents representing the state?

5. For a paternalistic policy to be successful, the state should possess gigantic amounts of information (Rizzo and Whitman, 2009). In particular, the state should know the following: 1) the 'true' preferences of individuals (a point to which we will return later); 2) which specific errors are made by individuals and in which specific situations; 3) what is the cost of these errors (e.g., the extent of their negative impact on welfare); 4) whether individuals take any steps themselves to prevent their errors and, if they do, to what extent such steps are effective; 5) how different errors are inter-related; 6) what is the effectiveness of possible corrective measures available to the paternalistic state?

It is clear that simply stating that people are subject to certain behavioral errors is not enough. The state should know the exact situations in which such errors are made. Without this knowledge, interventions by the state will appear to be either excessively extensive (involving situations wherein individuals do not make errors) or too narrow (neglecting certain situations wherein individuals make errors). The state should also understand the consequences of various behavioral anomalies in terms of welfare, and its interventions should be sized according to the severity of the particular error. Errors leading to huge losses in welfare would justify intervention, whereas errors leading to minor losses would not.

Individuals do not remain indifferent to their behavioral errors. Quite often they realize them and make focused attempts to eliminate them (for example, people prone to obesity or drunkenness can avoid places, situations and circumstances in which temptation may be too strong; they can train their will power, seek the support of people around them, etc.). If the state neglects all of these factors, instead of enhancing welfare, its interventions could make an outcome worse (for example, by strengthening external controls over human behavior, correspondingly weakening incentives towards self-control).

Different cognitive and behavioral errors may be inter-connected; they can either increase or neutralize each other's effects. Under such potentially contradictory conditions, it is extremely difficult to determine the optimum extent of intervention. Behaviors that seem to be optimal if only one particular anomaly is taken into account can turn out to be sub-optimal if all errors are taken into account simultaneously. In the absence of detailed information about how different biases are inter-connected, a paternalistic policy is likely to be counterproductive. For example, by correcting the optimism bias, the negative consequences related to the pessimism bias will increase.

Finally, the state should be capable of accurately estimating the effectiveness of various corrective measures(e.g., it when will be enough to provide more complete information, when it is more appropriate to manipulate frames, in what situations it is more effective to use 'sin taxes,' and in what situations are direct bans indispensable?)

Meeting these informational requirements is a practically irresolvable task. The volume of information needed by a paternalistic state to pursue a fine-tuning policy is hardly smaller than the volume of information that was needed by a central planning system for its effective functioning.

6. Behavioral economics criticizes the neoclassical concept of a hyper-rational individual while, in fact, promoting the state as a hyper-rational agent at the same time. This oversight is even more surprising as the multiple selves of an agent represented by the 'state' can be referenced, not figuratively, but quite literally.

Behaviorist reasoning implies that participants in the political process (voters, parties, experts, politicians and public officials) have an innate immunity to cognitive and behavioral errors. The reality is far from it. Governments are liable to errors as much as individuals. They have short planning horizons. They apply ultra-high discount rates when benefits from certain decisions can be obtained right away whereas the costs will be incurred in the distant future. They easily give in to temptations promising short-term political dividends. They tend to make promises and then break them when it is time to fulfill what has been promised. They often make decisions in a 'hot' psychological condition. And, as historical experience shows, most real political leaders can hardly be regarded as an ideal of clear reasoning and emotional balance (Glaeser, 2004).

Furthermore, there is every reason to believe that, in politics, behavioral anomalies are much more widely spread than in any other sphere. Compared to other spheres, in politics, the incentives to correct errors and to control irrational impulses on the parts of the policy makers are weaker as the consequences of the errors made by politicians affect them very little and primarily influence the society at large (Glaeser, 2004).

There is another factor to consider, the pressures applied by special interest groups. As a consequence, in political markets, risks of the exploitation of behavioral errors (manipulation of errors in the interests of individual politicians or lobby groups) appears to be much higher than in economic markets characterized by a tough competitive discipline.

Finally, it is important that the errors of individuals and states have entirely different consequences on welfare: 'Human beings, as rightfully noted by E. Glaeser, err, but the welfare losses created by private errors are surely second-order relative to the welfare losses created by governments. Individuals may procrastinate and invest foolishly, but they do not voluntarily enroll in concentration camps' (Glaeser, 2004. P. 412).

Seeking to restrict behavioral biases in private life, the behaviorist normative framework drastically extends possibilities for behavioral biases in politics. Hence, in many cases, it may happen that the 'medicine' appears to be much worse than the 'disease.'

7. Another critically important question is to what extent the results of laboratory experiments can be translated into human behavior 'in the field?' In the laboratory, subjects often find themselves in unusual conditions, must face problems that are new to them, and, in most cases, get merely a symbolic remuneration for participating in the experiment. In real life, people have the ability to gain experience and the profits and losses that reward them for rational decisions or punish them for irrational ones are incomparably larger. Thus, they have strong incentives to specialize and learn to avoid errors, incentives that are generally not found in laboratories.

Behaviorists tend to ignore the obvious fact that rational behavior is not an innate characteristic of man. It is gradually formed through the process of learning, and it is impossible to acquire it by skipping a (short or long) period of erroneous (irrational) decisions, during which experience is accumulated. People learn from errors and, in this context, rationality and behavioral biases are not necessarily mutually exclusive as is, in fact, suggested by behavioral economics. Among other things, this implies that in the attempt to protect individuals from demonstrating irrationality, a state's paternalistic policy may hinder individuals from gaining experience (i.e., developing rational behavioral skills), rendering them ultimately less rational.

The laboratory environment is institutionally sterile. It almost completely lacks the various tools for preventing and correcting behavioral errors that are available to people in real life. The behavior of a given person in institutionally poor and institutionally rich environments can be dramatically different. Suggestions that institutions (market ones, in the first place)

make human behavior more rational than it is by itself are a long-standing topic of economic theory, starting with Adam Smith. An effective system of institutions radically lowers the requirement for the rationality of economic agents and can make their behavior rational, even when it deals with agents suffering serious cognitive restrictions and/or lacking self-control. The available data confirm that the accumulation of experience, use of stronger incentives, access to various corrective mechanisms and exposure to the discipline of market competition reduces the frequency of behavioral errors quite sharply (see, for example: Levitt and List, 2007).

8. All of the foregoing discussions have proceeded from the implicit assumption that the 'new' paternalism offers a normative standard that allows for the consequences of certain behavioral errors to be assessed, both in relation to an individual and to the whole society. Let us recall that it is about the preferences reflecting the 'true' interests of people. However, when dealing with this fundamentally important issue, the 'new' paternalism appears to be impotent.

On the one hand, it claims to seek to improve the life of people based on their own preferences and asserts that the state knows what these 'true' preferences are better than the people themselves. On the other hand, it professes that individuals do not have any structured preferences, even formally, as evidenced by the incoherence of the peoples' choices. As a matter of fact, according to behavioral economics, any person simultaneously has several mutually exclusive sets of preferences.

This poses the question: can any of these sets be recognized as 'true' (that is reflecting a person's real interests) and, if yes, which one? Whether the behaviorist normative approach makes any sense at all depends upon the answer to this question. Indeed, if people do not have any 'true' preferences, then it becomes impossible to improve their lives according to such non-existing preferences!

It is here that behavioral economics demonstrates its evident conceptual arbitrariness. It offers no criteria for determining one set of preferences as 'true' and another as 'false.' If a coherence of preferences is the only condition for ensuring behavioral rationality, it can be achieved in many ways by picking certain consistent sub-sets out of all preferences and putting off all others. Which of the elements of this totality should then be sacrificed as not reflecting an individual's 'true' interests?

For example, in terms of hyperbolic discounting, coherence of preferences can be achieved either by reducing the short-term preference time rate to a long-term rate level (which behaviorists advocate) or by increasing the long-term rate to the short-term rate level. In cases where an individual's behavior is strongly influenced by the emotional background, coherence can be achieved by blocking decisions made in a 'hot' psychological state (the point insisted on by behaviorists) as well as decisions made in a 'cold' state. In the case of old-age savings, such coherence can be achieved by introducing the 'default' option of automatically including employees in retirement savings plans (as behaviorists demand) or not including them. Any of these ways to harmonize preferences is as good as the others. So, in what direction should individuals be 'nudged'? In the absence of clearly defined criteria, we cannot answer this question.

When behaviorists discern the true character of a person in his economical and not extravagant self; in his patient and not impatient self; in his reasonable and not impulsive self; in his 'cold' and not 'hot' self, etc., they do it in an arbitrary manner, proceeding solely from their own subjective likes and dislikes. As a result, the 'new' paternalism, which promised to respect the preferences of its 'paternalized' individuals, emerges into the classical 'old-fashioned' paternalism forcing its 'patron's' attitudes and values on individuals.

9. By proclaiming the endogeneity of preferences as a decisive argument against traditional welfare economics, the behaviorist approach, in fact, shakes its normative foundation. But, in this way, it destroys this foundation also for itself as the idea that individuals have a certain sub-set of 'true' preferences appears to be left hanging in the breeze. What are the possible ways of breaking this deadlock?

It would seem that one of them is empirical. If it were possible to demonstrate that agents with rational (coherent) preferences are much more successful and prosper more than agents with irrational (incoherent) preferences, it could serve as a criterion for determining 'true' preferences by clearing them from various behavioral errors. However, behaviorists are not willing to take this path. This reluctance may be related to the fact that there are practically no data suggesting that individuals, whose behaviors are based on the canons of the rational choice model, are more successful than individuals with noticeable behavioral deviations from such canonical behaviors. 'No studies we are aware of, observe A. Berg and G. Gigerenzer, show that those deviating from rational choice earn less money, live shorter lives, or are less happy' (Berg and Gigerenzer, 2010. P. 148).

Adherents of behavioral economics prefer a hypothetical approach to the empirical one. The hypothetical approach involves reconstructing 'true' preferences with the help of an as-if experiment. As we have already mentioned in reference to Sunstein and Thaler (see above), this approach suggests that replacing the factual preferences of individuals with the counterfactual preferences that would be demonstrated by individuals in ideal informational and cognitive conditions. The essence of such a reconstruction is best explained in the words of J. Harsanyi: 'A person's true preferences are the preferences he would have if he had all the relevant factual information, always reasoned with the greatest possible care, and were in a state of mind most conducive to rational choice' (Harsanyi, 1982. P. 55).

Then, 'true' preferences are preferences that would emerge in an individual's choices if the individual were an unboundedly rational being, i.e., had perfect information, unlimited cognitive capabilities and absolute will power. However, this is a rather strange solution. In fact, it suggests replacing ordinary people with some supernatural beings that cannot be found in real life.

First, strictly speaking, it is outside human knowledge to know what preferences such hyper-rational agents would have. One can only make more or less plausible speculations about it (Grüne-Yanoff, 2012). Second, it is not clear why the preferences of such agents should be completely coherent - as, just like ordinary people, they could be subject to status-quo bias (McQuillin and Sugden, 2012). Third, the very notion of 'complete information' cannot be regarded as unambiguously defined. Its interpretation depends upon the objectives set by an individual, his capabilities, and on the context in which he acts (McQuillin and Sugden, 2012). Fourth, endowing an individual with complete information, unlimited cognitive capabilities and absolute will power would change his personality so much that he would, in fact, be a totally different person. Can there be any respect for individual preferences in this case?

'New' paternalism is based on an overly ambitious and clearly unrealistic view of human nature; no actual person can comply with it by definition. This clearly shows that, in its normative constructions, behavioral economics applies to principles of human behavior that are external to it, which runs contrary to its promise to judge welfare based exclusively on a person's own aims and aspirations: 'the consequence of dividing the self has been to undermine the very idea of true preferences. If true preferences don't exist, the libertarian paternalist cannot help people get what they truly want. He can only make like an old fashioned paternalist, and give people what they should want' (Leonard, 2008. P. 259).

Because the notions of the hypothetical preferences of imaginary hyper-rational beings proposed by 'new' paternalism advocates as a normative ideal inevitably remain hazy and vague, the resulting void is filled with quite specific notions held by various groups of agents. These may be the preferences of incumbent politicians, or their expert consultants, the public majority, etc. A closer look at which of the alternative sets of preferences the 'new' paternalists agree to be 'rational' clearly reveals that they always reflect the values of the middle class, i.e., the social stratum to which the experts themselves belong. They invariably prefer economy over extravagance, caution over impulse, calculation over emotions, a healthy lifestyle over bad habits, patience over impatience, long-term planning over momentary plea-

sures, and so on. It can be said that from a sociological point of view, 'new' paternalism is none other than an attempt to dictate middle class values to society. At this point, the differences between the 'new' and the traditional 'old' paternalism are purely nominal. At the end of the day, both try to dictate alien behavioral norms to people against their will.

10. As we have noted, the main target of the behaviorist attack is the orthodox approach of welfare economics with its normative standard of satisfying individual preferences. However, social philosophy and economic theory have an alternative normative tradition that has not fallen under the above criticism.

This tradition does not trust ordinary people to make the best decisions all of the time. It only argues that we do not know who could make better decisions than people themselves and that politicians, officials and experts are almost sure to do a worse job. It does not postulate that people are infallible hyper-rational beings. On the contrary, they are viewed as boundedly rational and boundedly ethical, shortsighted, weak-willed beings prone to temptation. As for the 'consumer sovereignty' principle, it is defended not because 'the consumer is always right,' but because it holds freedom of choice itself as a high value. It puts a person's right to err among the essential human rights because it is a necessary condition for their autonomy, and helps them to become more competent, responsible and independent (in other words, more rational).

We are speaking about the tradition rooted in the teachings of Adam Smith and other philosophers of the Scottish Enlightenment, which can be broadly termed as the tradition of classical liberalism. According to F. Hayek, this tradition sees a human 'not as a highly rational and intelligent but as a very irrational and fallible being, whose individual errors are corrected only in the course of a social process' (Hayek, 1948. P. 9). Its originators never viewed human behavior as optimizing: '...in their view man was by nature lazy and indolent, improvident and wasteful, and ...it was only by the force of circumstances that he could be made to behave economically or carefully to adjust his means to his ends' (Hayek, 1948. P. 11).

While behavioral economics derives paternalist conclusions from bounded rationality, the Smithian tradition comes to a directly opposite, anti-paternalistic conclusion. It argues that in contemporary, complex societies, the main cognitive load in coordinating economic activity is borne by institutions of 'extended order' (term coined by Hayek) rather than individual minds (Boettke and Caceres, 2013). We can say that the main purpose of these institutions is to prevent and correct the errors made by boundedly rational individuals. If people were fully rational beings and possessed unlimited cognitive powers, such institutions would become redundant. Adherents of the behavioral approach do not realize that by using paternal 'care' to rationalize the behavior of individual agents, they are hampering the work of ' institutions of 'extended order', complicating economic coordination and rendering the whole process less rational. As a result, the measures they propose to cut the costs of behavioral errors may lead to cost increases, and society may begin shifting toward greater irrationality instead of moving toward the higher rationality for which they are arguing.

9. Conclusion

Behavioral economics is a new advanced field of study that has greatly altered modern economic theory. Although the majority of formal models used by economists are still built on the perfect rationality principle, behavioral economics has presented significant empirical evidence of how far real human behavior deviates from rational assumptions. It made a real breakthrough in normative economic analysis, discarding its traditional anti-paternalistic attitude.

However, contrary to first impression, the attitudes of behavioral economics towards the conventional rational model of choice turn out to be far from straightforward. Behavioral economics rejects it as a descriptive theory but indirectly retain it as a normative ideal, and movements toward it are viewed as advancing the ultimate good. Instituting more detailed

governmental 'care' for society is proclaimed to be the most effective means for achieving this goal. From this point of view, the situation with the perfect rationality concept strongly resembles situation with another basic idea of neoclassical economic theory — that of perfect competition. Modern economists are also inclined to reject the perfect competition model as an adequate description of economic reality, yet accept it as a normative standard, calling on state interventions to be guided by it.

Past experience shows that each new type of 'market failure' discovered by economists has been followed by an outbreak of state activity. Behavioral economics is not an exception. It has identified a number of behavioral 'market failures,' paving the way for a wide range of paternalistic interventions. Its adherents not only suggested a great number of ingenious schemes not previously used in state regulation but also attached intellectual respect to many traditional policies that had been used ad hoc, without a clear nod from economic theory.

This theoretical justification, while sanctioning certain types of state intervention, also makes it more disciplined because other types of intervention, to which such justifications turn out to be inapplicable, are beginning to be considered unacceptable. This is how matters previously stood with traditional welfare economics. Its theoretical arguments could be used successfully to justify various redistribution programs, as well as different measures aimed at overcoming external effects. At the same time, the same arguments provided grounds for normative prohibition of paternalistic interventions that directly interfered with individual decision-making.

This did not prevent the state from resorting to various paternalistic 'care' measures not sanctioned by economic theory. However, 'non-theoretical' forms of state intervention have a different status (i.e., that of inevitable evil or manifestation of the state's weakness) and provide more limited growth potential. It is one thing to raise minimum wages under pressure from trade unions, and quite another to raise them in accordance with the theory of monopsony to increase employment. It is one thing to have inflated a budget deficit during the pre-Keynesian era, but another to pursue the same policy based on Keynesian arguments about the necessity of aggregate demand stimulation. For a long time, the lack of a necessary theoretical sanction has been holding back the 'inherent' paternalistic ambitions of modern governments and preventing them from expanding their influence.

The situation changed when behavioral economics provided such a sanction. It demonstrated that, in matters of their own welfare, people are quite rarely the best judges. This led to the conclusion that a paternalistic state might considerably improve the quality of people's decision making. When the 'inherent' instincts of the state receive intellectual support from the academic community, it is most often a prologue to increased government activity. This, as far as one can judge, has happened with the ideas of behavioral economics and the normative program of 'new' paternalism built upon these ideas. In this context, some observers identify a growing trend for the welfare state to transform into a paternalistic state with much broader opportunities and authority to oversee the activities of individuals.

Behavioral economics has criticized the orthodox approach of welfare economics, demonstrating the instability of individual preferences and their dependence upon the decision-making context. Paradoxical as it may be, this type of criticism turned out to be detrimental for its own political program. It is undermining the normative basis not only for the traditional welfare approach but also for the 'new' paternalism, which is proving to be powerless in solving its most important question regarding 'true' individual preferences. In these conditions, it would only be natural to turn to an alternative normative tradition in social philosophy and economic theory, i.e., a 'liberal' one in the broadest sense, built upon the idea of freedom, rather than the idea of welfare. It is of utmost importance that, unlike 'new' paternalism, in this alternative tradition, associated with the names of Hume, Smith, Mill, Knight, Hayek, and Coase, the bounded rationality of an individual is not seen as an argument for expanding but rather for limiting state intervention in the economy and, more broadly, in people's private life.

References

Angner E., & Loewenstein, G. (2007). Behavioral economics. In U. Maki (Ed.), *Handbook of the philosophy of science, Vol. 13: Philosophy of economics* (pp. 641–689). Amsterdam: Elsevier.

Ariely, D. (2008). Predictably irrational. Harper: HarperCollins Publishers.

Arrow, K J. (1996). Preface. In J. K. Arrow, E. Colombatto, & M. Perlman (Eds.), *The rational foundations of economic behaviour* (pp. xiii-xvii). London: Macmillan.

Basham, P. (2010). *Are nudging and shoving good for public health?* Unpublished manuscript, Democracy Institute.

Berg, N., & Gigerenzer, G. (2010). As-if behavioral economics: Neoclassical economics in disguise? *History of Economic Ideas, 18* (2), 133–166.

Boettke, P., Caceres, W. Z., & Martin, A. (2013). Error is obvious, coordination is the puzzle. In R. Frantz, & R. Leeson (Eds.), *Hayek and behavioral economics*. London: Palgrave Macmillan.

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for 'asymmetric paternalism'. *University of Pennsylvania Law Review, 151* (1), 1211–1254.

Choi, J., Laibson, D., Madrian, B. C., & Metrick, A. (2002). Defined contribution pensions: Plan rules, participant decisions, and the path of least resistance. In J. M. Poterba (Ed.), *Tax policy and the economy: Vol. 16* (pp. 67–113). Cambridge, MA: MIT Press.

Costa-Font, J. (2011). Behavioural welfare economics: Does 'behavioural optimality' matter? *CESifo Economic Studies, 57* (4), 551–559.

Glaeser, E. L. (2004). Psychology and the market. *American Economic Review, 94* (2), 408–413.

Grune-Yanoff, T. (2012). Old wine in new casks: Libertarian paternalism still violates liberal principles. *Social Choice and Welfare, 38* (4), 635–645.

Grunwald, M. (2009). How Obama is using the science of change. *Time Magazine,* April 2.

Harsanyi, J. (1982). Morality and the theory of rational behaviour. In A. Sen, & B. Williams (Eds.), *Utilitarianism and beyond* (pp. 39–62). Cambridge: Cambridge University Press.

Hayek, F. A. (1948). *Individualism and economic order*. Chicago: Chicago University Press.

Jolls, C., & Sunstein, C. R. (2006). Debiasing through law. *Journal of Legal Studies, 35* (1), 199–242.

Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185* (4157), 1124–1131.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47* (2), 263–291.

Katona, G. (1951). *Psychological analysis of economic behavior*. N. Y.: McGraw-Hill.

Klein, D. B. (2004). Statist quo bias. *Economic Journal Watch, 1* (2), 260–271.

Knetsch, J. L. (1989). The endowment effect and evidence of nonreversible indifference curves. *American Economic Review, 79* (5), 1277–1184.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112* (2), 443–477.

Lambert, C. (2006). The marketplace of perceptions. *Harvard Magazine, 2*, 50–95.

Leonard, T. C. (2008). Review of "Richard H. Thaler, Cass R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness". *Constitutional Political Economy, 19* (4), 356–360.

Levine, D. K. (2012). *Is behavioral economics doomed? The ordinary versus the extraordinary.* Cambridge: Open Book Publishers.

Levitt, S. D., & List, J. A. (2007). Viewpoint: On the generalizability of lab behavior to the field. *Canadian Journal of Economics, 40* (2), 347–370.

Loewenstein, G., & Haisley, E. (2006). The economist as therapist: Methodological ramifications of 'light' paternalism. In A. Caplin, & A. Schotter (Eds.), *The handbook of economic methodologies: Vol. 1. Perspectives on the future of economics: Positive and normative foundations* (pp. 210–245). Oxford: Oxford University Press.

McQuillin, B., & Sugden, R. (2012). Reconciling normative and behavioural economics: The problem to be solved. *Social Choice and Welfare, 38* (4), 553–567.

Mitchell, G. (2005). Libertarian paternalism is an oxymoron. *Northwestern University Law Review, 99* (3), 1245–1277.

Plott Ch. R., & Zeiler, K. (2007). Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory? *American Economic Review, 97* (4), 1449–1466.

Plott Ch. R., & Zeiler, K. (2011). The willingness to pay-willingness to accept gap, the "endowment effect", subject misconceptions, and experimental procedures for eliciting valuations: Reply. *American Economic Review, 101* (4), 1449–1466.

Radnitzky, G., & Bernholtz, P. (Eds.) (1987). *Economic imperialism: The economic approach applied outside the field of economics.* N. Y.: Paragon House Publishers.

Rizzo, M. J., & Whitman, D. G. (2009). The knowledge problem of the new paternalism. *Brigham Young University Law Review, 4*, 905–968.

Saint-Paul, J. (2011). *The tyranny of utility. Behavioral social science and the role of paternalism*. Princeton: Princeton University Press.

Sent, E.-M. (2004). Behavioral economics: How psychology made its (limited) way back into economics. *History of Political Economy, 36* (4), 735–760.

Simon, H. A. (1987). Bounded rationality. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new Palgrave.* N.Y.: W.W. Norton.

Sugden, R. (2008). Why incoherent preferences do not justify paternalism. *Constitutional Political Economy, 19* (3), 226–248.

Sunstein, C., & Thaler, R. (2003a). Libertarian paternalism. *American Economic Review, 93* (2), 175–179.

Sunstein, C., & Thaler, R. (2003b). Libertarian paternalism is not an oxymoron. *University of Chicago Law Review, 70* (4), 1159–1202.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization, 1* (1), 39–60.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science, 4* (1), 199–214.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth and Happiness*. New Haven & London: Yale University Press.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59* (4, Pt. 2), S251–S278.

Waldfogel, J. (2005). Does consumer irrationality trump consumer sovereignty? *Review of Economics and Statistics, 87* (4), 691–696.

Weale, A. (1991). Paternalism. In D. Miller et al. (Eds.), *The Blackwell encyclopaedia of political though* (pp. 367–368). Oxford: Basil Blackwell.

Wintour, P. (2010). David Cameron's 'nudge unit' aims to improve economic behaviour. *The Guardian,* September 9.

Wright, D. J., & Ginsburg, D. H. (2012). Behavioral law and economics: Its origins, its fatal flaws, and its implications for liberty. *Northwestern University Law Review, 106* (3), 1–56.